Project Proposal

COSC 4P02 January 16th, 2024
Software Engineering II Brock University
Winter 2024

Definitions

Herein, and in future documents, the following definitions apply:

- Issue tracker: product backlog
- Issue: backlog item, or user story
- Tracking issue: epic
- Sprint goal: sprint backlog item

Team Structure

QiQi Gao Product Owner/Dev... qg17jh@brocku.ca 6416762	Michael Boulet Scrum Master/Dev... mb20ot@brocku.ca 7063548	Meet Patel Developer mp20dp@brocku.ca 7056708	Michael Noyes Developer mn17bg@brocku.ca 6261374
Hamza Yousuf Developer hs19uo@brocku.ca 6772149	Jonathan Coletti Developer jc21jq@brocku.ca 7333339	Jose Henriquez Developer jh20uo@brocku.ca 7088792	David Fawzy Developer df20ft@brocku.ca 7084593

Proposal

Problem

Web Summarizer and Shortener.

Objectives

- Use NLP techniques and LLMs to return a brief summary of a webpage.
    - Web Page provided through a single URL.
    - Scraping web pages
    - Transcription and summarization of videos
    - Custom summarization level (e.g. sentence level, paragraph level, to a tenth grader, etc.)
- Allow authenticated users to track their summarized pages and shortened URLs
- Develop an interactive dashboard to summarize webpages, login, and manage links.
- Develop a web extension to quickly summarize and link to web pages.

Importance

- Let users quickly grasp the essence of long web pages & videos, cutting out the clutter
- Enhances productivity and learning efficiency

Software Engineering Process

Github will be our main code and scrum collaboration tool, responsible for hosting our code, issue tracker, pull requests, code review, and CI. Our team's public Github organization can be found here. Hosted therein is:

- the main project repository,
- the main project's issue tracker,
- the document/report tracking repository, and,
- the scrum tracking project.

Depending on the results of the requirements process, and what software architecture, language, and frameworks are selected as a result, the following may be implemented depending on feasibility and time constraints: code style guides, continuous integration, a production environment, automatic deployment, package managers, and additional architecture specific tools. These processes will be planned if and when their need arises.

Commiting Workflow

We will aim to have git branches follow a master/develop branches in combination with feature branches. The master branch will track the latest stable release, while develop tracks current development. Feature branches may branch off of develop and rebase back into develop when progress finishes. One-off changes may be directly committed to development. The master

branch will not be directly committed to, and will be rebased from development when a new release is ready.

Feature branches and one-off commits can be reviewed by developers via Github's by creating a pull request. Other developers will then review it and either approve the pull request or request changes. A certain number of approvals can be set as required for a pull request to be merged.

Scrum

Our scrum sprint schedule will start on January 23rd, with each sprint lasting for 2 weeks, until either: the project concludes, or the term ends, whichever occurs first. Team sprint goals will be tracked on the respective github project on a per-sprint basis. Individual developer's goals for the sprint will be self-selected on every sprint planning session under the guidance of the product owner, and the current state of the issue tracker.

Recurring progress meetings should be held weekly every Tuesday at 17:00 UTC-05:00 (time subject to change). On alternating progress meetings that coincide with a new sprint, the next sprint will be planned. Communication, including meetings, will be performed over Discord.

Issue Tracker

The issue tracker will be populated as the project progresses, with issues being added as they are identified. Issues may be tagged with the appropriate tags that identify & classify the issue. Features, other functionality, or tracking issues may be added to the issue tracker as they arise, with or without discussion. Tracking issues should either: identify issue numbers of their related issues, or preferably, enumerate all relevant issues under the tracking issue as subgoals. Finally, sprint goals may be converted to and linked to an issue using the appropriate github tool when appropriate.

After issues are identified on the tracker, on subsequent sprints, they may be worked on depending on their severity and relevance to the current sprint goals. Alternatively, if an issue arises that is found to block a current sprint goal's progress, work on the issue may be done immediately and should be noted in the sprint goal's description.

Contribution Tracking

Contributions will be mainly tracked by git commit history, and pull request/issue contributions. In the case that multiple authors worked on a commit, a co-author should be added to the commit. Attendance to meetings will be automatically tracked out-of-band by a Discord bot. Message history in discord channels will also be used as contribution tracking when deemed necessary.